Indian Journal of Economics & Business, Vol. 4, No. 1, (2005) : 73-94

Asset Pricing Simultaneity, Three-Factor Model and Cost Analysis

Robert D. Coleman

Abstract[1]

The aim is to explain the phenomenon of the atheoretic size and value effects in asset pricing, which lack face validity, and to thereby resolve contradictory assertions about the theoretical and empirical validity of asset-pricing models of return with price-entailing risk factors. It is shown that size, value and other price-entailing factors are not scientifically valid in an asset-pricing model of return estimated and tested by scientific statistical methods. Analysis shows that the direct cost of these pricing fallacies is hundreds of millions dollars each year in capital markets worldwide.

Key Words : Capital asset pricing, portfolio, multifactor model, size, book-to-market value ratio, earnings-price ratio, simultaneity, logical circularity, fallacy, scientific validity, cost-benefit analysis.

JEL Classification : G12 - Asset Pricing, C12 - Hypothesis Testing.

Introduction

There is a contradiction about the validity of a class of risk factors specified in multifactor asset-pricing models. This contradiction has not been resolved in the literature.

Ball (1978) presents earnings' yield (E/P) and dividend yield (D/P) as asset-pricing risk factors in the role of proxies. Ironically Ball undermines his argument by mentioning the non-independence of these factors and the entire class of yield factors in models of return. Fama and French (FF) in FF (1992) cite Ball but are silent on his disclaimer about non-independence. Fogler (1982) emphasizes the need to insure uncorrelated residuals of sample securities, meaning individual securities rather than portfolios of securities. Ironically he lists capitalization size among his approved risk factors.

Hall and Hall (1993, p. 12), referring to isolation of the left-hand side (LHS) of an equation, asserts: "By analogy, in a standard regression setting, one can never include as a right-hand variable a variable that is an important part of the left-

Robert D. Coleman, 3715 Bryn Mawr Drive, Dallas, Texas 75225.
E-mail : rcoleman @ mba1971.hbs.edu

hand variable." Heaton (1993, pp. 36-37), a comment on Hall and Hall (1993), says: "There has been renewed focus on variables such as book-to-market value, firm size, price-earnings ratios, and dividend-price ratios in explaining differences in returns. Hall and Hall dismiss such results because they often use the stock price as part of the explanatory variables." Additional comments by panelist N. Gregory Mankiw and the other participants—Olivier Blanchard, William Brainard, Richard Cooper, Bronwyn Hall, Robert Hall, William Nordhaus, Andrei Shleifer and James Tobin—appear to agree with the assertion by Hall and Hall, but it is not supported by argument, demonstration, analysis or proof.

In contradiction to Hall and Hall, Schwert (1983, pp. 9-10), an introductory summary of a symposium on the size factor in multifactor financial asset-pricing models, asserts:

> A search for an explanation of this [size effect] anomaly has been unsuccessful.… None of the attempts to modify the CAPM … have been successful at discovering the 'missing factor' for which size is a proxy. …I hope these empirical regularities will stimulate future research on the aspects of capital market institutions that will explain what now seem to be anomalies. New models of asset pricing that can explain the empirical evidence on the 'size effect,' while maintaining the assumption of rational maximizing behavior, would be a significant step forward in financial research (single-quotes in the original).

Schwert as well as FF (1992, 1993) assert the scientific validity of price-entailing risk factors such as size specified in an asset-pricing model of price-entailing return. This assertion directly contradicts the Hall and Hall assertion, clarified by Heaton, that in effect price-entailing risk factors such as size can not be specified validly in a model of price-entailing return. This contradiction is neither necessary nor sufficient to our purpose but helps to motivate the discussion. The existence of logical circularity due to econometric simultaneity, I believe, resolves their differences and puts the "size effect" and related fallacies to rest, not through mere assertions but rather through argument and evidence.

Fama-French Asset-Pricing Models

The most prominent supporters of the size effect include Fama and French. In contrast to the theory-based univariate CAPM, FF (1993) presents an atheoretical five-factor asset-pricing model (FF5F) with two bond and three stock risk factors. Fama, French, Booth and Sinquefield (1993) presents the resulting Fama-French Three-Factor Model (FF3F) with risk factors based on market, size and distress. Distress, measured by book-to-market value (B/M), is an alleged proxy for low or negative earnings. FF (1992) credits Banz (1981) for the size effect, and it credits Stattman (1980) and Rosenberg, Reid and Lanstein (1985) for the book-to-price effect. Reinganum (1979a, 1979b, 1981) are alleged to confirm the market value effect or size effect of Banz (1978, 1979, 1981).

In FF (1993) the equity risk factor *SMB* (small minus big) is based on size, and the equity risk factor *HML* (high minus low) is based on value. *SMB* and *HML* are zero-net-investment arbitrage portfolios. Size is measured by market capitalization or market value of equity, defined as share price multiplied by number of shares outstanding. Value is measured by B/M, defined as book value of equity divided by market value of equity. B/M is alleged to explain more than does size [FF (1992, p. 440)]. B/M is book value/size and is used to distinguish value stocks from other stocks. In contrast, the price-earnings ratio (P/E) is used to distinguish growth stocks from other stocks. Fama and French mention other equity risk factors: E/P ratio, annual earnings per share divided by share price; leverage, debt divided by size; and dividend yield, annual dividends per share divided by share price. Size, B/M, E/P, leverage and dividend yield are price-entailing variables, and each entails price as a non-trivial component. FF3F is alleged to outperform the CAPM in terms of R^2, explanatory power.

A CAPM or a multifactor model may incur simultaneity in estimation of the model if the stock on the LHS is included in the index portfolio used as a proxy for the market in the calculation of market-beta or the market risk factor. This can be avoided by omitting the stock from the index portfolio when the stock appears on the LHS. The discussion of simultaneity, therefore, focuses hereinafter on the FF3F size and value risk factors.

FF (1993, p. 37), Table 9a (iii), shows intercepts for 25 stock portfolios formed on size and B/M, July 1963 to December 1991, 342 months, for the Fama-French Two-Factor model (FF2F):

$$R(t) - RF(t) = a + sSMB(t) + hHML(t) + e(t)$$

In FF (1993) and in other FF articles not cited here because they are redundant negative examples, Fama and French make nine mistakes, and each mistake invalidates their asset pricing model. The mistakes in FF5F, FF3F and FF2F are as follows.

1. Failure to show *a priori* reasons or theoretical grounds for their price-entailing risk factors, which lack face validity.
2. Failure to identify the price-entailing variables in their asset pricing model of return and to clearly and fully disclose the resulting simultaneity in their model equation.
3. Failure to specify a statistical non-deterministic non-definitional relationship.
4. Failure to isolate the LHS in their empirical asset-pricing model equation.
5. Failure to use a full-information simultaneous-equation system or a limited-information single-equation recursive system to estimate and test their model.
6. Failure to adhere to standard linear regression theory for credible RHS factors.
7. Failure to report results of standard specification error tests for exogeneity, covariance equations for correlated variables, and uncorrelated-residuals tests

for the given sample. The 1963-1990 monthly data do not represent the widest range of market behavior, and the sample is too small for acceptable correlation tests.

8. Failure to fully disclose *ad hoc* test methods. The sample is split by unstated procedures. Partition of stocks into a fewer number of portfolios results in loss of effective data observations and, if nonrandom, adds information.

9. Failure to use standard diagnostic tests and, instead, design non-independent, *ad hoc*, *ex post facto*, diagnostic tests of induced explanatory power. Ancillary mistakes include using mimicking returns for variables that entail price, and pattern recognition that is subjective, retrospective and thus not reliable.

Before a simultaneous equations system can be estimated, the equations must be specified. The non-existence of information about these fundamental macroeconomic and financial relationships is due to the non-existence of the associated reverse proxies. The estimation of such a simultaneous equations system therefore is not possible. In short, the FF models are logically circular due to embedded simultaneities and thus are not scientifically valid.

The conclusion of this first argument is that Fama and French with FF3F violate the crucial factor-independence assumption of classical linear regression theory and thus deviate from standard scientific methodological practice. FF3F and the related asset pricing models, therefore, are logically circular due to risk factors with unresolved embedded direct theoretical and empirical simultaneity. As a result, size, B/M and other price-entailing risk factors are not scientifically valid in a single-equation asset pricing model of return estimated and tested by standard scientific linear regression methods.

FF3F *Ad Hoc* Diagnostic Test

Concerning the split-sample *ad hoc* diagnostic test of FF3F for factor independence, FF (1993, p. 47), Fama and French say:

> Many readers worry that the apparent explanatory power of *SMB* and *HML* is spurious, induced by the regression setup. We think this is unlikely. … Still, an independent test is of interest.

Ironically the FF *ad hoc* diagnostic test of FF3F is not independent and not scientifically valid for two reasons, one empirical and the other theoretical.

A. Portfolio Formation Factors

The empirical reason why the split-sample test, or split-portfolio test in terms of asset allocation, is not valid is because the splitting of the sample or master portfolio into each of the 25 sub-portfolios formed on size and on B/M is a distraction. FF do not say how they split each portfolio into two equal groups. In addition the number of portfolios is arbitrary and has no economic justification. But the important fact to pay attention to is that the formation of portfolios in effect uses new factor information

to sort the data *before* the sample is split. This factor information is indirectly built into the structure of the model in lieu of specifying factors directly in the model.

An equivalent equation for model estimating and testing specifies a dummy or qualitative variable for portfolios with 25 categories, one for each portfolio. To avoid the dummy-variable trap, the situation of perfect multicollinearity, only 24 dummy variables are specified. Usually the lowest or the highest category is not specified, and the lowest often makes more sense. For example the lowest of the 25 portfolios is formed by the intersection of the lowest size quintile and the lowest book-to-market value quintile.

Through portfolio formation, therefore, the pricing model surreptitiously incorporates the two variables on which the portfolios are formed, size and book-to-market value. As a result, any subsequent splitting of the sample portfolio to achieve independence is illusory, and the split-portfolio test is not independent. Rather the split-sample test makes the same mistake as the specified model that it is designed to diagnose. The split-sample test includes an irresolvable embedded indirect simultaneity, a logical circularity.

B. Deterministic Versus Statistical

The theoretical reason why such a contrived model-specific test for induced explanatory power of risk factors is not reliable is the nature of the equation to be solved. An equation that is true by definition has perfectly known relationships among its terms, variables and components. It is a functional or deterministic identity and a mathematical certainty. In contrast, an equation for which the correct variables and the correct relationships among the variables are unknown is not an identity; and when the variables are uncertain or probabilistic, it is a stochastic model of random variables.

An equation that includes elements of definitions, not-uncertain deterministic relationships and uncertain statistical hypothesis, such as a simultaneity that is not modeled by a simultaneous-equation system, can be referred to as a hybrid equation. FF3F is not a purely statistical relationship as required for standard linear regression models. FF3F is a hybrid. It is a combination of a definition with formulaic certainty and a stochastic regression model of a statistical hypothesis. The simultaneity, however embedded, is still logically circular. The solution of such a hybrid equation is confounded due to inexplicable results that depend on the model, the sample and any data adjustments.

The distinction between deterministic and statistical relationships cannot be overemphasized. It is also important not to confuse the distinction between deterministic and statistical relationships with the distinction between the deterministic and stochastic parts of an equation. The hybrid equation is a combination of definition, deterministic relationship and statistical hypothesis, whereas a linear regression equation is a combination of a deterministic part and a stochastic part. The deterministic part consists of the explanatory variables, given

or known, and fixed or constant in repeated samples. The stochastic part is the error term [Maddala (1992, pp. 62-64)].

The conclusion of this second argument is that Fama and French, with their split-sample *ad hoc* diagnostic test of FF3F, for the second time violate the crucial factor- independence assumption of linear regression theory and thus deviate from standard scientific methodological practice. The FF *ad hoc* diagnostic test is logically circular due to risk factors with unresolved embedded indirect simultaneity. Size, book-to-market value of equity (B/M) and other price-entailing risk factors, when specified as portfolio formation factors, are not scientifically valid in a single-equation asset pricing model of return estimated and tested by standard scientific linear regression methods.

Common Mistakes

High school algebra teaches us to isolate the LHS of an equation in order to avoid logical circularity and simultaneity. Our arguments address the same fundamental mistakes in FF3F and in the split-sample *ad hoc* diagnostic test. Both involve bad procedures and bad reporting, but not necessarily bad data. In essence, FF3F and the *ad hoc* diagnostic test of FF3F are both aspects of the same irremediable bias-inducing fatal fallacy.

Fama and French react to the criticism of FF3F as it is presented by FF in FF (1992). Their reply appears in FF (1993), and they do not recant. FF implicitly acknowledge the need for independence of the risk factors, and they respond with their split-sample *ad hoc* diagnostic test of factor independence. The *ad hoc* test of the spurious induced explanatory power of the non-independent risk factors is, in itself, not independent. Rather, it allows for surreptitious influence of the risk factors. The convoluted *ad hoc* diagnostic test is designed specifically for the FF3F and is not one of the diagnostic tests found in the standard scientific methods used by researchers in economics.

In the same article [FF (1993)] in which FF attempt to prove, at least for their sample, the independence of the atheoretic risk factors, using their model-specific *ad hoc* diagnostic test, they also create an orthogonalized version of their market risk factor in order to achieve maximum independence relative to other variables in the stock-pricing model. FF thus demonstrate a high degree of knowledge about factor independence and great sophistication in designing spurious non-scientific tests of factor independence.

A. Standard Diagnostic Tests

The logically prior theoretical question is: In any statistical model are the explanatory variables and their entailed components independent of the dependent variable and its components in order to isolate the LHS? The theoretical answer is a quantitative equation based on analysis of the variables irrespective of any data sample. The logically subsequent empirical question is: If not, are the entailed

components a significant part of the LHS and fatal fallacies (see Appendix)? The empirical answer is quantitative and sample-specific.

Next in logical order is the empirical question: In any linear regression model are the explanatory variables independent of the error term? The direct empirical answer is qualitative — either yes or no — depending, among other things, on the logical theoretical independence of the explanatory and dependent variables. If X or a component of X is entailed in Y, then X cannot be logically independent of the error term, $Y - Y_c$. The quantitative answer is indirect and empirical, provided by standard diagnostic tests. Such tests of any regression model are made using the full unsorted sample. Thus such tests of FF3F are made on the full sample of stocks not sorted and partitioned into portfolios.

The most common statistical test of simultaneity is a version of Hausman's specification test [Hausman (1978)], a general and widely used procedure for testing the hypothesis of no misspecification in the model. In order to use the OLS principle to estimate a model we specify that each explanatory factor is independent of the error term, i.e., it is exogenous. If an explanatory factor is not correlated with the error term, i.e., is exogenous or predetermined, we are justified in estimating the equation by OLS. In adapting Hausman's test to exogeneity, we can regard the single-equation simultaneity model as one equation in a simultaneous-equation model [Maddala (1992, pp. 506-507)].

In addition thorough examination of residuals involves graphical or numerical analysis of the residuals for any unusual patterns that might indicate violation of the crucial assumptions of the classical linear regression model (CLRM). The plots of residuals versus each explanatory variable, the dependent variable, the predicted values and the observation number may be inspected for this purpose. Furthermore, in the special case of a definitional deterministic model equation such as FF3F, we can apply the equation for covariance between two variables that are correlated to demonstrate the existence of uncorrelated error terms. Also FF3F reformulated to eliminate simultaneity is not linear but rather is a quadratic relationship.

FF are silent on the standard scientific diagnostic tests and do not explain why these tests results are not disclosed in their article. Instead FF (1993, p. 14) reports the simple correlation between each of the three equity risk factors with one another for the average sample return, but not for individual stock return, for a time-series of 342 monthly observations, but not a cross-section or a panel of firm-specific observations. The test data reported by FF are incomplete, inadequate and not full disclosure.

The further conclusion to the first two arguments is that Fama and French, by failing to perform and report standard diagnostic tests of the model specification for risk factor independence, for the third time violate the crucial factor-independence assumption of the CLRM and thus deviate from standard scientific methodological practice. The FF asset pricing models are logically circular due to direct simultaneity, and the split-sample *ad hoc* test of FF3F is logically circular due to indirect

simultaneity. The CLRM allows no exception for any possible sample size. FF3F and the split-sample *ad hoc* diagnostic test, therefore, are irredeemably invalid. As a result size, B/M and other price-entailing risk factors are not scientifically valid in a single-equation asset-pricing model of return estimated and tested by the CLRM and the Gauss-Markov Theorem.

Validity Versus Interpretation

Berk (1995) diverts attention from the basic question of non-validity of factors to the superficial question of interpretation of factors as proxies of some unknown variables. Ball (1978) asserts that the price-earnings ratio is a proxy for unnamed variables. FF (1992, p. 428) cites Ball's proxy argument, and Fama, French, Booth and Sinquefield (1993) asserts that size and B/M are proxies for the financial distress of a firm. Berk thus repeats the same mistake made by Ball, Schwert, Fama, French and others.

These authors ignore the essential nature of a proxy variable and get it backwards. The logical direction is from a variable that is known but unobservable or difficult to measure, say UNOBS, to a proxy or instrumental variable that is known, observable and more easily measurable, say OBS. Going logically from UNOBS to OBS is like a trustee who votes in the role of proxy for a *known* shareholder, but going illogically from OBS to UNOBS is like a trustee who votes as a proxy for an *unknown* nonexistent shareholder. Not only is the backward direction from OBS to UNOBS illogical, it is also wasteful. A backward proxy results in loss of information, degradation of data and mere speculation. Such a reverse proxy is a violation of true method because it goes from the lesser known to the better known, whereas a valid proxy goes from the better known to the lesser known.

At the empirical testing level Berk misrepresents a valid proxy and ignores the crucial requirement for no correlation between the proxy and the error term as explained in Gujarati (1988, p. 609). The logically circular FF3F models of unresolved undisclosed simultaneity in articles by Berk and others are scientifically neither interesting nor important. Fama and French co-authored five more articles on FF3F after FF (1993), yet the FF3F literature speculates on factor interpretation while overlooking the logically prior issue of factor validity.

The conclusion to the third argument is that Fama and French, in their interpretation of their atheoretic risk factors as proxies for some unspecified unknown variables, for the fourth time violate the crucial factor-independence assumption of the CLRM and thus deviate from standard scientific methodological practice. The FF asset pricing models, therefore, are logically circular due to risk factors with unresolved embedded direct simultaneity. The diversionary FF3F controversy about risk factor interpretation masks the underlying FF3F fatal fallacy based on simultaneity, and it displaces scientifically valid risk factors such as the CAPM market-beta.

FF (1992, p. 428) asserts that "Ball's proxy argument for E/P might also apply to size (ME), leverage [L/ME] and book-to-market equity [BE/ME]" because according

to Keim (1988) "all these variables can be regarded as different ways to scale stock prices." Yet scaling share price violates the isolation of the LHS. Reinforcing the conclusion to the third argument, Fama and French, in their interpretation of the atheoretic size and value risk factors as different ways to scale stock prices, for the fifth time violate the crucial factor-independence assumption of linear regression theory and thus deviate from standard scientific methodological practice. The FF asset pricing models, therefore, again are logically circular due to risk factors with unresolved embedded direct simultaneity.

Invariance to Sample

The international evidence for size and value [FF (1998)] is not surprising, since the mathematical definition of return is invariant to time, place, currency, bourse and marketable security. FF3F or any other econometric model is mathematically invariant across all empirical samples that are randomly drawn and not sorted according to model-specific logically circular variables, and thus it applies to all time zones, nations, languages, currencies, publicly traded companies, stocks, stock markets and bourses. Empirical results in one time, place, situation and circumstances, therefore, are likely to replicate empirical results in any other time, place, situation and circumstances.

It also is not surprising that a logically circular model that has statistically significant results for one sample of a given number of observations turns out to have significant results for any other sample of equal or larger size. All of the empirical data are used to estimate the same model; and if the model is fatally flawed due to logical circularity, then the data sampling and data quality are irrelevant.

The conclusion to the fourth argument is that Fama and French, in the estimation of FF3F with diverse data samples across nations, for the fifth time violate the crucial factor-independence assumption of the CLRM and thus deviate from standard scientific methodological practice.

Invariance to Methodology

The Maroney and Protopapadakis (2002) [MP (2002)] abstract says:

> The positive relation of returns with Book-to-Market ratio (BE/ME) and their negative relation with Market Value (MVE) remains strong under a general stochastic discount function (SDF) that does not depend on a specific asset pricing model and avoids potentially serious simultaneity biases inherent in the Fama and French three-factor model.

The phrase "potentially serious" is an understatement – it should say "fatal" instead, because logically circular price-entailing explanatory variables are not unavoidable biases that do not qualitatively change statistical significance of tests of the model, rather they are avoidable fatal fallacies. MP (2002) is scientifically

neither interesting nor important. In addition, the result is not surprising because logical circularity is invariant to any scientific statistical model in any field of inquiry. Once again, the relation of return with either BE/ME or MVE is always logically circular because these relations are true by definition. They are not stochastic relations. This is basic logic and common sense.

The size and book-to-market value factors are atheoretic, and there are no accepted fundamental determinants associated with them as stated in several published articles, including MP (2002, p. 190), footnote 2: "There have been several but as yet not wholly successful attempts to find a theoretical explanation for this empirical regularity." Size and book-to-market value interpretations are discussed in MP (2002, pp. 189-190).

The BE/ME and MVE explanatory variables have a prior and thus over-riding relationship with return that is logically circular independent of any specified model. This logical relationship supersedes any given particular model specification that includes price-entailing risk factors to explain price-entailing return, whether the model specifies a general stochastic discount function or any other theoretical or empirical relationship.

MP (2002, p.190) says: "It is at best doubtful that these tests ['of the significance of the *BE/ME* and *MVE* effects reported to-date'] satisfy the OLS requirement that explanatory variables are independent of the LHS variable." It says on page 191: "Our tests of the *BE/ME* and *MVE* effects are "spanning" tests that use Stochastic Discount Factors or SDFs. … The tests consist of constructing candidate SDFs from a subset of portfolios (reference portfolios) and using these to price or "span" all the assets in the market." It says on page 191, footnote 5: "Including the largest and smallest ranked *BE/ME* or *MVE* portfolios is equivalent to including the *HML* [Fama-French (1993) High Minus Low *BE/ME*] portfolio in a regression." It says on page 193: "Equally-weighted portfolios for each country are formed at the end of each October of year *t* by sorting stocks by Market Value of Equity (*MVE*) and Book-to-Market Equity (*BE/ME*) separately (for one-way sorts), and by sorting first by *MVE* and then by *BE/ME* for two-way sorts." It says on page 193, footnote 10: "These definitions [*BE/ME* and Book Value (BE)] are consistent with Fama and French (1992)."

In MP (2002) the "subset of portfolios (reference portfolios)" is equivalent in terms of logical circularity to one of the two sub-samples of the FF (1993) split-sample *ad hoc* diagnostic test of independence of the explanatory risk factors in the FF3F model. Thus these "spanning" tests likewise entail embedded indirect simultaneities that manipulate the data by sorting it and forming groups or portfolios on logically circular price-entailing variables, *BE/ME* and *MVE*, with either a one-way sort using one of the variables or a two-way sort using both of the variables. Indirect simultaneity due to sorting the data on a logically circular variable, for both FF (1993) and MP (2002), is econometrically equivalent to direct specification of the logically circular variable in the model. This renders the model and its testing

and estimation logically circular and thus fallacious, meaningless, indeterminate and scientifically not valid.

It is not true, therefore, that a general stochastic discount function independent of a specific asset pricing model avoids the simultaneity inherent in the FF approach. It may avoid the direct simultaneity inherent in FF (1992) due to model specification, but it does not avoid the indirect simultaneity inherent in FF (1993) due to data sorting. In addition the avoidable simultaneities are fatal fallacies. MP (2002) cites FF (1992) but does not cite FF (1993). The MP (2002) "spanning" tests, just like the FF (1993) split-sample *ad hoc* diagnostic test, entail an unresolved embedded indirect simultaneity by data sorting on logically circular variables. Thus the MP (2002) model and argument fail to circumvent over-riding logical circularity in the form of either fatal direct simultaneity or fatal indirect simultaneity.

FF3F clearly does not meet the methodology requirements of Ordinary Least Squares (OLS) estimation. And changing from OLS to the Generalized Method of Moments (GMM) estimation methodology does not escape the logical circularity in the FF3F model. OLS, GMM and all other scientific statistical methodologies do not affect the logical circularity of the underlying specified model that is being tested and estimated. A price-entailing explanatory variable, such as size or book-to-market value, is logically circular in a model of price-entailing return, and this logical circularity in FF3F or any other econometric model is invariant across all scientific methods of testing and estimation.

The conclusion to the fifth argument is that FF3F is invariant in terms of logical circularity across all scientific methods of model testing and estimation, and for the sixth time violates the over-riding logically prior laws of reasoning by entailing the fallacy of logical circularity in the form of an embedded direct or indirect single-equation simultaneity and thus deviates from standard scientific methodological practice.

It has been shown in the first part of this paper that the arguments in support of the FF3F model are specious. They either subvert the CLRM behind the Gauss-Markov Theorem which is a fundamental result for econometricians or subvert basic scientific logic. In particular, size and value are not independent [Gujarati (1988, pp. 52, 56-57), Pindyck and Rubinfeld (1991, pp. 95-97)]. In addition FF3F and related atheoretic asset-pricing factors are not scientifically valid, but rather are logically fallacious. None of the subsequent papers published by Fama and French address these issues or correct the fallacies pertaining to FF3F in FF (1992) and the split-sample *ad hoc* diagnostic test of FF3F in FF (1993).

FF3F Cost Analysis

A thorough detailed analysis of the costs of the FF3F model and its atheoretic risk factors, collectively referred to as FF3F, is beyond the scope of this paper. For purposes of scholarship or science, the theoretical arguments are sufficient. The magnitude of the impact of the FF3F fallacy, nevertheless, is of interest to those

who may be adversely impacted. Those so impacted include investors, financial practitioners and others outside the academic community.

Any stock may be selected for rational reasons, even if it also would be selected due to irrational fallacious factors. But if the stock is selected because of a fallacious factor, and would not be selected without the fallacious factor, the investor incurs additional costs attributable solely to the fallacious factor.

Stock selection factors are important. Reports of finding anomalous risk factors that create arbitrage opportunities for superior risk-adjusted expected returns relative to the broad market averages have received widespread acceptance since the size effect was published [Banz (1981)]. For example small-cap value-style stocks, relative to large-cap growth-style stocks, are asserted to offer premium risk-adjusted expected returns and superior long-term investment performance. There is no scientific reason why FF3F-based investment strategies could consistently earn superior long-term results; nor is there any scientific reason why the FF3F-contrarian large-cap (e.g., S&P 500 Index) and growth (e.g., MSCI EAFE Index) strategies could do so either.

Following such misguidance many investors worldwide have invested large amounts of money for more than 20 years based on the fallacies of size, book-to-market value and earnings-price ratio as factors in stock screening and selection. This is only part of the impact in the stock market alone, and the impact goes well beyond the stock markets and bourses of the world. Stock trading to rebalance portfolios to mimic logically circular risk factors is counterproductive, avoidably materially economically wasteful, costly and tax-inefficient. Sharpe (1991, p. 7) says:

If "active" and "passive" management styles are defined in sensible ways, it *must* be the case that

1. before costs, the return on the average actively managed dollar will equal the return on the average passively managed dollar and
2. after costs, the return on the average actively managed dollar will be less than the return on the average passively managed dollar.

These assertions will hold for *any* time period. Moreover, they depend *only* on the laws of addition, subtraction, multiplication and division. Nothing else is required. [Italics are in the original.]

Passive management is defined as investing in the market portfolio in market proportions, and active management is anything that is not passive.

A. Costs

Beyond academia the costs of FF3F are vast, widespread and long-running. FF3F, like any capital asset-pricing or equity-pricing model adopted by practitioners, is crucial to three important applications: first, price of capital is cost of capital, and price of equity is cost of equity; second, price of capital is fair rate of return for regulatory purposes; and third, price of capital is discount rate in valuations of

companies. Specific examples abound with evidence of continuing harm to the academic, practitioner, rate-regulation, valuation, taxation, legal, and investment communities. The following cost analysis will focus on the harm of FF3F to stock investors.

In the stock exchanges and bourses throughout the world FF3F has caused an aggregate massive transfer of wealth from large-cap to small-cap stocks and from growth-style to value-style stocks. Millions of investors worldwide with billions of dollars invested in stocks, both directly in public markets and indirectly through employer-sponsored retirement-savings plans, have been adversely impacted because of transaction costs and asset-management fees regardless of the price performance of the underlying marketable securities. Although the total effect on any particular investor is difficult to determine, Arrow (1964) demonstrates that the aggregate effect on all investors in the stock market is a misallocation of societal risk-bearing. In the case of FF3F this is due to a shift from large-cap and growth-style stocks to small-cap and value-style stocks.

One of the pioneers and leaders in offering and selling financial products explicitly based on the FF3F model and its logically circular risk factors is Dimensional Fund Advisors Inc (DFA), a family of equity index mutual funds. DFA is a small part of the worldwide financial services industry that profits from application of the fallacious size and value factors in the design of financial products. Scores of mutual funds intentionally promote and invest in small-cap, large-cap, value-style (high B/M) and growth-style (low B/M) stocks. The following are presented in chronological order of fund inception. AXA Rosenberg is another fund family focused on the quantitative size and value factors. Fidelity Investments is the largest mutual fund company. As of December 31, 2002, Fidelity total assets in eight stock funds based on size (small-, mid- and large-cap) or B/M (value- and growth-style) are $7.1 billion, which is 1.6% of $432.8 billion in 242 equity funds and 1% of $713.1 billion in mutual fund assets under management, which is about 11% of the $6.6 trillion U.S. mutual fund industry. Vanguard Group is the second largest mutual fund company and has the largest fund. TIAA-CREF is the largest fund company intended for investors who are teachers and professors.

A comparison of the expense ratios shows that the stock funds structured on size (small-, mid-, and large-cap) and value or B/M (value-, blend-, and growth-style) have a cost range of 8 to 245 basis points. The midrange or median fee markup is 69 basis points for size and value and 106 basis points for enhanced size and value. Each fee markup is the excess above total stock market indexes, across U.S. and non-U.S. stocks and across institutional and individual investors, as shown in Table 1.

To estimate total annual cost at each mutual fund, we need to accumulate daily cost at each mutual fund, and this requires data for the daily assets and the daily excess fee markup. Since we have incomplete information to make such a calculation, we can use the unweighted average excess fee markup for our small sample of the largest mutual funds based on size and value. For U.S. size and value stock funds

Table 1.A: Size and Value Funds: Expense Comparisons

	Expense Ratio[1]				
Fund Family	**DFA**	**AXA Rosenberg**	**Fidelity**	**Vanguard**	**TIAA-CREF**
Fund Assets	$35 Billion	$16 Billion	$713 Billion	$560 Billion	$270 Billion
Fund Inception:					
Size (Market Equity)	1981	1989	1994	1997	2002
Value (Bk-to-Mkt Equity)	1993	1997	2001	1997	2002
US – Individual:					
Total Stock Market	N/A	N/A	41 – 45	20	N/A
S&P 500[2]	30 – 39	115	41	12 – 18	43
Size & Value[3]	31 – 56	140 – 159	78 – 119	17 – 124	30 – 44
Size & Value Enhanced[4]	34 – 76	242 – 248	124 – 175	17	N/A
US – Institutional:					
Total Stock Market	N/A	N/A	41 – 45	6 – 8	N/A
S&P 500[2]	16	75	39	2	8 – 14
Size & Value[3]	17 – 111	100 – 115	78 – 119	10	8
Size & Value Enhanced[4]	N/A	195 – 218	124 – 175	10	14
Non-US – Individual:					
Total Stock Market	N/A	N/A	N/A	37	N/A
MSCI EAFE[5]	N/A	135	57	34	49
Size & Value[3]	34 – 156	178	180	34 – 67	49
Size & Value Enhanced[4]	76	N/A	N/A	31	N/A
Non-US – Institutional:					
Total Stock Market	N/A	N/A	N/A	N/A	N/A
MSCI EAFE[5]	N/A	160	57	21	15 – 20
Size & Value[3]	102	150	180	N/A	15
Size & Value Enhanced[4]	164	N/A	N/A	20	N/A

Contd...

Table 1.A : (Contd...)

Combination of Funds and Fund Families					
Fund Type	**Inception**	**Range**	**Midpoint**	**Markup bp**	**Markup %**
Total Stock Market	1999	6 – 45	25	N/A	N/A
S&P 500[2]	1976	2 – 115	58	33	132%
MSCI EAFE[5]	1981	15 – 180	92	67	268%
Size & Value[3]	1981	8 – 180	94	69	276%
Size & Value Enhanced[4]	1996	17 – 245	131	106	424%

Notes :

1. Ratio of total annual operating expense to average net assets, low-high range in basis points (1 basis point or bp = 1/100 of 1%).
2. Benchmark is Standard & Poor's 500 Composite Stock Index or comparable index.
3. Size as market equity capitalization (small-, mid- and large-cap), and Value as book-to-market value of equity (value- and growth-style).
4. Size and Value with active trading to increase returns above benchmarks or to increase tax-efficiency.
5. Benchmark is Morgan Stanley Capital International Europe, Australasia, Far East Index or comparable index.

N/A is Not Applicable.

Table 1.B: Size And Value Stock Mutual Funds

Small-Cap Value: A convenience sample of stock mutual funds based on their names:	
Berger Small Cap Value Fund	Quintana Small Cap Value Fund
Eclipse Small Cap Value	T Rowe Price Small Cap Value
GE Small-Cap Value Equity Fund	Third Avenue Small-Cap Value Fund
Northern Small Cap Value Fund	Tocqueville Small Cap Value Fund
Pioneer Small Cap Value A	Turner Small Cap Value Fund
Large-Cap Growth: A contrarian investment strategy exemplified by stock mutual fund names:	
Eaton Vance Large-Cap Growth Fund	Munder Large-Cap Growth Fund

sold to individual investors, the average excess fee markup for size and value is 80 basis points, and the average excess fee markup for enhanced size and value is 117 basis points. The average excess fee markups for U.S. size and value stock funds sold to institutional investors and for non-U.S. size and value stock funds sold to individual and to institutional investors are significantly higher. Our estimates of the minimum total cost are based on a global weighted-average excess fee markup of 20 basis points, much lower than the 80 basis points and 117 basis points of the lowest-cost category for size and value in our sample.

The size and value factors are continuous variables but have become best known through different versions of the familiar three-by-three category table as shown in Table 2. Many of the leading stock equity indexes such as the S&P 500 have size, value-style and growth-style components, and they are used as performance benchmarks for mutual funds.

Table 2.A: Size and Value 3-by-3 Table

	Value	**Blend**	**Growth**
Large	Large Cap Value	Large Cap Blend	Large Cap Growth
	NEXT BEST		**WORST**
Middle	Mid Cap Value	Mid Cap Blend	Mid Cap Growth
	Small Cap Value	Small Cap Blend	Small Cap Growth
Small	**BEST**		**NEXT WORST**

Note : The FF3F model factors based on value, size and market can be represented in an X-Y-Z three-dimensional space. According to the FF3F model, at any market level, on average over the long-term on a risk-adjusted basis the small-cap value-style stock returns are allegedly superior to large-cap growth-style stock returns, and value (book-to-market value of equity ratio) explains more than does size (market equity capitalization). Size is inversely related to return, and value is directly related to return. Low value is growth-style; high value is value-style. The third factor is based on market, a proxy for the whole stock market, in the Z dimension not shown. Size and value are by definition mathematically related to return; thus they are logically circular and their explanatory power is spuriously induced. The unsurprising results follow automatically due to simultaneity, to information in each variable and to multivariate correlation signs.

Table 2.B. Applications of the Size and Value 3-by-3 Table

1. AXA Rosenberg Fund Analysis with Market Cap by Value/Growth Style.
2. Dimensional Fund Advisers Inc. (DFA) three-dimensional presentation of FF3F model by its dimensions of Market, Size and Value.
3. Fama-French Three-Factor (FF3F) model with Market, Size and Value factors.
4. Fidelity Investments StyleMap with Size and Value/Growth factors.
5. Financial Engines Asset Allocation with Size and Geographical Location asset classes, and Mutual Fund Styles with investment strategy categories such as Value, Growth and Blend.
6. Index Fund Advisers Inc. (IFA) three-dimensional presentation of FF3F model by its dimensions of Market, Size and Value.
7. Morgan Stanley Equity Investment Style with Size and Value/Growth factors.
8. Morningstar Style Box with Size and Value/Growth factors.
9. Vanguard Group Tic-Tac-Toe board with Size and Value/Growth Style factors.
10. Equity indexes subdivided according to Size and Value/Growth Style factors (e.g., CRSP, Dow Jones, Lipper, MSCI, Russell, S&P and Wilshire).

Conclusions

What is here found true for size is also true for value (B/M) and other price-entailing risk factors in return models. For brevity of discussion, size can be used as a representative of the class of price-entailing risk factors and hereinafter refers to the whole class.

Size when specified in an asset-pricing model of return is logically circular due to irresolvable embedded simultaneity. Price-entailing size is not independent of price-entailing return. Thus size is neither a theoretically valid nor an empirically valid risk factor to explain expected return. Size specified directly in a return model is a fatal fallacy and not a priced risk factor. FF3F is logically circular in the form of direct simultaneity due to model specification, and thus FF3F is fallacious, meaningless, indeterminate and not scientifically valid. In short FF3F is not scientifically interesting or important.

Likewise size as a portfolio formation factor is not valid with a return model. The FF split-sample *ad hoc* diagnostic test of FF3F factor independence is logically circular in the form of indirect simultaneity due to data sorting and manipulation. Thus the FF split-sample *ad hoc* test of the FF3F factors is not independent of FF3F itself. Size used as a portfolio formation variable to sort data to test and estimate a return model is a fallacy.

Furthermore size as a proxy factor directly specified in a return model is an erroneous use of the concept of a proxy; it is illogical, wasteful, backwards and violates method. Size specified as such a proxy factor is not scientifically valid but rather is a mistake. Also size specified as a scaling factor is an erroneous use of scaling variables. Size specified as such a scaling variable is not scientifically valid but rather is a mistake.

These research findings are invariant to scientific research methodology. They are also invariant to sample and apply to all times, places, currencies, bourses and marketable securities. The findings are based on mathematics and scientific logic, and they are independent of situation and circumstances.

Any representation, therefore, that scientific studies show that investment strategies and financial products based on either the FF3F model or its risk factors can earn consistent average risk-adjusted expected returns greater than the returns on broad market indexes or conventional market proxies is not true. There can be no possible evidence or proof that FF3F is scientifically valid. FF3F has never been scientifically valid and never will be as long as logically circular econometric simultaneity is not scientifically valid.

The cost of FF3F since 1981 is widespread and has continued for over 20 years. The enormous harmful costs impact several communities worldwide including investors, academic research, finance education, finance practitioners, rate-regulated industries and taxpayers. A large class of investors includes employees participating in retirement-savings plans, and another class includes high-net-worth families.

The cost of FF3F that is easiest to quantify for stock investors is excess management fees charged to stock index mutual fund investors, but these are only part of the total costs. Based on ultra-conservative estimates of only transaction costs as reflected in excess fund-management fees, without including market price distortions, inefficient risk-bearing allocations and intangible human costs, the logically circular FF3F model and related logically circular risk-factors cause economic waste that amounts to hundreds of millions of U.S. dollars each year and more than a billion U.S. dollars since inception.

Direct costs to investors alone, estimated using mutual fund financial report data, realistically could be at least double these ultra-conservative estimates and are more likely to be quadruple these estimates. Investment strategies and financial products including stock mutual funds, stock indexes and stock index mutual funds that are explicitly designed to exploit the spurious superior risk-return characteristics of size and value are scientifically meaningless and economically wasteful.

It appears, therefore, that the Fama-French Three-Factor Model is the biggest fallacy in stock market history when measured by several criteria: the number of investors adversely impacted, the total direct cash cost to all investors, the total harm to all persons adversely impacted, the geographical range of residence of the adversely impacted persons, the duration since the first commercial application of the FF3F model or its price-entailing risk factors, and the number of professionals in academia and the financial-services industry who are involved.

Appendix: Fatal Fallacies and Biases

A fatal fallacy is not merely a source of bias. It is irremediable and thus ends an argument. Most empirical articles include a section for diagnostics in which the specified models are subjected to various standard diagnostic tests including tests for biases attributable to known sources. For example if variables on both sides of a model equation are adjusted with a price-level index, this would introduce logical circularity in the form of single-equation simultaneity and thus could induce spurious significance in model test results. Note that the price-level index is an adjustment applied to the variables rather than a separately specified variable. The model can be tested with the adjustment for price-level inflation included in one case and excluded in another case to demonstrate that the results are not qualitatively different between the cases. If there is no qualitative difference in outcome then the adjusted model is unnecessary, but it may be reported to present the model variables in a familiar or desired form. If there is a qualitative difference in outcome, then the source of bias should be eliminated from the model or the model be rejected. The burden of proof is on any researcher who publishes a known biased model and argues that the bias is not significant.

The most meaningful distinction and criterion for the degree of bias is significant versus non-significant, i.e., whether or not the known bias qualitatively changes the outcomes of estimating the model. The two qualitative outcomes are: reject, or fail to reject. A null hypothesis, e.g., that there is no relation between each of the

explanatory variables and the dependent variable, can fail to be rejected, but strictly speaking no null hypothesis is said to be accepted. If a model estimate is statistically significant at conventional levels of probability, then the model and its null hypothesis are rejected; and if an estimate is not significant, then the model and hypothesis fail to be rejected.

All known sources of bias should be disclosed by the author. If an avoidable source of bias does not qualitatively change the outcome of model estimation, then it should not be specified in the model for two reasons: first, it is avoidable bias; and second, it is too small to make a qualitative difference in outcome and thus is superfluous. If an avoidable source of bias does qualitatively change the outcome of model estimation, then it should not be specified in the model for two reasons: first, it is avoidable bias; and second, it is large enough to make a qualitative difference in outcome and thus is a fatal fallacy. For bias resulting from known avoidable sources there are thus no exceptions at the empirical level just as there are no exceptions at the theoretical level.

Logical circularity is a known avoidable source of bias. Price-entailing risk factors specified in a model of price-entailing return lead to a fatal fallacy and thus should be eliminated from the model. Simple comparative tests cannot be made of the FF3F model because its simultaneity involves a non-detachable part of the definition of the explanatory variables themselves rather than a detachable adjustment to the variables. Nevertheless valid metrics for size of firm such as total sales revenue and for financial distress such as debt-to-book equity ratio could be specified in the adjusted model, and the FF3F model with its logically circular explanatory variables would be the unadjusted model. When the adjusted and unadjusted FF3F models are compared in this manner, there is usually a qualitative difference in results for all scientifically valid metrics of size and of distress and for all random samples not sorted by logically circular variables.

True scientific methods are not fallacious or biased; rather, specified models and data manipulation may be fallacious or biased. Certain kinds of biases are unavoidable, but known logical circularity and single-equation simultaneity are not unavoidable. Rather they usually can be avoided in legitimate empirical studies, and otherwise they can be fully disclosed and tested for exceptions such as the price-level indexing of variables on both sides of the model equation. The phrases "logical-circularity bias" and "simultaneity bias" are misleading if not inherently meaningless when they refer to such an avoidable fatal fallacy. They are symptomatic of the minimization characteristic of psychological denial when uncritically adopting certain intra-disciplinary traditions such as the Fama-French Three-Factor Model in financial economics.

The scientific way to treat known avoidable simultaneity is to summarily reject the model and not obfuscate it with diagnostic testing for bias. A fatal fallacy cannot be corrected; that is why it is called fatal. Diagnostic testing of logically circular statistical models such as the FF3F model is analogous to an eye doctor performing diagnostic vision tests for astigmatism on a congenitally blind patient who can never

use his eyes for vision. Astigmatic bias can be treated with corrective lenses to remove the distortion, but vision is fatally flawed or totally impaired by congenital blindness.

The detection of undisclosed fatal fallacies and the diagnostic testing of known biases are not performed in a vacuum but rather in light of what decisions are impacted and thus of what is at stake for those who believe, act on and trust the research findings. The Fama-French Three-Factor Model and its price-entailing factors have materially adversely impacted millions of investors alone worldwide for more than two decades.

NOTE

1. I would like to thank Dr. Irving Hoch, University of Texas at Dallas, for his comments on a late-stage draft. Also appreciated are the comments of anonymous referees. The usual disclaimer applies.

REFERENCES

Arrow, K.J. (1964), "The Role of Securities in the Optimal Allocation of Risk-Bearing", *Review of Economic Studies,* Vol. 31, pp. 91-96.

Ball, R. (1978), "Anomalies in Relationships between Securities' Yields and Yield-Surrogates", *Journal of Financial Economics,* Vol. 6, pp. 103-126.

Banz, R.W. (1981), "The Relationship between Return and Market Value of Common Stocks", *Journal of Financial Economics,* Vol. 9, No.1, pp. 3-18.

Banz, R.W. (1979), "On the Relationship between the Market Value and the Return of Common Stocks", *Proceedings: Seminar on the Analysis of Security Prices*, Center for Research in Security Prices (Sponsored by Merrill Lynch, Pierce, Fenner & Smith Inc), Graduate School of Business, University of Chicago, Vol. 24, No. 1, pp. 131-158.

Banz, R.W. (1978), *Limited Diversification and Market Equilibrium: An Empirical Analysis,* PhD dissertation, University of Chicago. http:// wwwlib.umi.com / dxweb / details ?doc_no=2034150.

Berk, J.B. (1995), "A Critique of Size-Related Anomalies", *Review of Financial Studies*, Vol. 8, No. 2, pp. 275-286.

Fama, E.F. and K.R. French (1998), "Value versus Growth: The International Evidence", *Journal of Finance,* Vol. 53, No. 6, 1975-1999.

Fama, E.F. and K.R. French (1993), "Common Risk Factors in the Returns of Stocks and Bonds", *Journal of Financial Economics,* Vol. 33, No. 1, pp. 3-56.

Fama, E.F. and K.R. French (1992), "The Cross-Section of Expected Stock Returns", *Journal of Finance*, Vol. 47, No. 2, pp. 427-465.

Fama, E.F, K.R. French, D.G. Booth and R.A. Sinquefield (1993), "Differences in the Risks and Returns of NYSE and NASD Stocks", *Financial Analysts Journal,* Vol. 49, No. 1, pp. 37-41.

Fogler, H.R. (1982), "Common Sense on CAPM, APT, and Correlated Residuals", *The Journal of Portfolio Management*, Vol. 8, No. 4, pp. 20-28.

Gujarati, D.N. (1988), *Basic Econometrics*, 2/e, McGraw-Hill, New York.

Hall, B.H. and R.E. Hall (1993), "The Value and Performance of U.S. Corporations", *Brookings Papers on Economic Activity,* Vol. 1, pp. 1-49.

Hausman, J.A. (1978), "Specification Tests in Econometrics", *Econometrica,* Vol. 46, No. 6, pp. 1251-1271.

Heaton, J. (1993), "Comments and Discussion", in "The Value and Performance of U.S. Corporations", *Brookings Papers on Economic Activity,* Vol. 1, pp. 35-46.

Keim, D.B. (1988), "Stock Market Regularities: A Synthesis of the Evidence and Explanations", in Elroy Dimson, ed.: *Stock Market Anomalies*, Cambridge University Press, Cambridge.

Maddala, G.S. (1992), *Introduction to Econometrics*, 2/e, Macmillan, New York.

Maroney, N. and A. Protopapadakis (2002), "The Book-to-Market and Size Effects in a General Asset Pricing Model: Evidence from Seven National Markets", *European Finance Review*, Vol. 6, No. 2, pp. 189-221.

Pindyck, R.S. and D.L. Rubinfeld (1991), *Econometric Models and Economic Forecasts*, 3/e, McGraw-Hill, New York.

Reinganum, M.R. (1981), "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings' Yields and Market Values", *Journal of Financial Economics*, Vol. 9, No. 1, pp. 19-46.

Reinganum, M.R. (1979a), "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings' Yields and Market Values", *Proceedings: Seminar on the Analysis of Security Prices*, Center for Research in Security Prices (Sponsored by Merrill Lynch, Pierce, Fenner & Smith Inc), Graduate School of Business, University of Chicago, Vol. 24, No. 2, pp. 225-282.

Reinganum, M.R. (1979b), *Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings' Yields and Market Values*, PhD dissertation, University of Chicago. http:// wwwlib.umi.com / dxweb / details ?doc_no=1529270.

Rosenberg, B., K. Reid and R. Lanstein (1985), "Persuasive Evidence of Market Inefficiency", *Journal of Portfolio Management*, Vol. 11, No. 2, pp. 9-16.

Schwert, G.W. (1983), "Size and Stock Returns, and Other Empirical Regularities", *Journal of Financial Economics,* Vol. 12, No. 1, pp. 3-12.

Sharpe, W.F. (1991), "The Arithmetic of Active Management", *Financial Analysts Journal,* Vol. 47, No. 1, pp. 7-9.

Stattman, D.W. (1980), "Book Values and Stock Returns", *The Chicago MBA: A Journal of Selected Papers,* Vol. 4, pp. 25-45.